Norwegian University of Science and Technology Department of Mathematical Sciences	TMA4180 Optimisation I Spring 2017 **Solutions to exercise set 3**

1 **a)** $\nabla f = Qx - b$ and $\nabla^2 f = Q$ from calculus. Since Q is symmetric positive definite (SPD), it follows that f is strictly convex on $\mathbb{R}^n$, and as such, there is at most one global minimum of f. Furthermore, this global minimum x^* must be a stationary point satisfying $\nabla f(x^*) = 0$. We conclude that $x^* = Q^{-1}b$, since Q is invertible (all eigenvalues of Q are positive, and hence, different from zero).

b) Newton's method reads

$$x_{k+1} = x_k + p_k,$$

where $p_k = -(\nabla^2 f_k)^{-1}\nabla f_k$, with $\nabla f_k := \nabla f(x_k)$, and $\nabla^2 f_k := \nabla^2 f(x_k)$. Therefore, given any x_0, we find that

$$x_1 = x_0 - Q^{-1}(Qx_0 - b) = Q^{-1}b = x^*,$$

that is, Newton's method converges after one step.

c) With chosen direction p_k in position x_k, and selected preliminary step length α and some parameter $c_1 \in (0,1)$, backtracking line search iteratively reduces α until

$$f(x_k + \alpha p_k) \le f(x_k) + c_1 \alpha \nabla f_k^\top p_k. \qquad (\star)$$

In our case,

$$f(x_k + \alpha p_k) = f(x_k) + \alpha \nabla f_k^\top p_k + \tfrac{1}{2}\alpha^2 p_k^\top \nabla^2 f_k \, p_k$$

by exact Taylor expansion (f is a quadratic form), and

$$p_k^\top \nabla^2 f_k \, p_k = -\nabla f_k^\top (\nabla^2 f_k)^{-\top} \nabla^2 f_k \, p_k = -\nabla f_k^\top p_k,$$

because $\nabla^2 f$ (and its inverse) is symmetric. Inserting these expressions into ($\star$) yields that

$$\left(1 - \frac{\alpha}{2}\right) \alpha \nabla f_k^\top p_k \le c_1 \alpha \nabla f_k^\top p_k.$$

Now, since p_k is a descent direction (Q is SPD), this gives the criterion

$$c_1 \le 1 - \frac{\alpha}{2}.$$

Hence, if $\alpha = 1$, then $c_1 \le 1/2$ is required, and, conversely, if $c_1 \le 1/2$, then the method accepts step lengths $\alpha \le 2(1 - c_1)$; in particular, $\alpha = 1$ is OK.

2 **a)** $\nabla f = (4x - 2y + 6x^2 + 4x^3, 2y - 2x)$ and

$$\nabla^2 f = \begin{bmatrix} 4 + 12x + 12x^2 & -2 \\ -2 & 2 \end{bmatrix},$$

Hence, stationary points satisfy $y = x$ by the first component of ∇f, while the second component yields that $0 = 2x(1 + 3x + 2x^2) = x(x + 1)(2x + 1)$. Thus critical points of f are $(0, 0)$, $(-\frac{1}{2}, -\frac{1}{2})$, and $(-1, -1)$. Now,

$$\nabla^2 f(0,0) = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} = \nabla^2 f(-1,-1) \qquad \text{and} \qquad \nabla^2 f(-\tfrac{1}{2}, -\tfrac{1}{2}) = \begin{bmatrix} 1 & -2 \\ -2 & 2 \end{bmatrix}$$

has eigenvalues $3 \pm \sqrt{5} > 0$ and $(3 \pm \sqrt{17})/2$ (one positive, and one negative), respectively. We conclude that $(0, 0)$, and $(-1, -1)$ are strict local minima, while $(-\frac{1}{2}, -\frac{1}{2})$ is a saddle point. Moreover, since $\nabla^2 f$ remains SPD both for $x > 0$ and $x < -1$ (the value of y is irrelevant), it follows that $(0, 0)$ and $(-1, -1)$ are the only candidates for global minima. Evaluating $f(0, 0) = 0 = f(-1, -1)$, shows that both are global minimisers of f.

b) Gradient descent method equals $(x_{k+1}, y_{k+1}) = (x_k, y_k) + p_k$, with $p_k = -\nabla f_k$. Starting with preliminary step length α, $\rho = 1/2$, and $c_1 = 1/4$, we accept a new step provided

$$f\big((x_0, y_0) + \alpha p_0\big) \leq f(x_0, y_0) + c\alpha \nabla f(x_0, y_0)^\top p_0 = 1 - 2\alpha$$

using that $p_0 = -\nabla f(x_0, y_0) = (2, -2)$.
Beginning with $\alpha = 1$, we reject the first try since $f\big((x_0, y_0) + \alpha p_0\big) = 13 > -1$. Reducing to $\alpha \leftarrow \rho\alpha = 1/2$, still gives rejection, but $\alpha = 1/4$ succeeds, because $f\big((x_0, y_0) + \alpha p_0\big) = 1/16 \leq 1/2$. Hence, we put $(x_1, y_1) = (-\frac{1}{2}, -\frac{1}{2})$,and proceed with a new round. However, (x_1, y_1) is a critical (saddle) point for f, so the gradient method stops here, thereby failing to converge to a minimiser.

c) Similarly as in b), the backtracking acceptance criterion for Newton's method reads

$$f\big((x_0, y_0) + \alpha p_0\big) \leq f(x_0, y_0) + c\alpha \nabla f(x_0, y_0)^\top p_0 = 1 - \tfrac{1}{2}\alpha,$$

since $p_0 = -\nabla^2 f(x_0, y_0)^{-1} \nabla f(x_0, y_0) = (0, -1)$ and $c_1 = 1/4$. Starting with $\alpha = 1$, we have $f\big((x_0, y_0) + \alpha p_0\big) = 0 \leq 1/2$, so the step is accepted. We then put $(x_1, y_1) = (x_0, y_0) + p_0 = (-1, -1)$. This point is a global minimiser, the conclusion being that Newton's method converged in one step.

3 See file tma4180s17_ex03_3.m on the website.